DOCUMENT RESUME

ED 324 950 FL 018 932

AUTHOR Moita Lopes, Luiz Paulo da
TITLE The Evaluation of an EFL Reading Syllabus Model: Integrating Experimental and Ethnographic/Qualitative Data.
PUB DATE Apr 90
NOTE 31p.; Paper presented at the Meeting of the World Congress of Applied Linguistics, sponsored by the International Association of Applied Linguistics (9th, Thessaloniki, Greece, April 15-21, 1990).
PUB TYPE Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Course Evaluation; *Curriculum Design; Educational Theories; *English (Second Language); Ethnography; Evaluation Methods; Foreign Countries; Models; *Reading Instruction; Second Language Instruction
IDENTIFIERS *Brazil (Rio de Janeiro); *Schema Theory

ABSTRACT

The research method used to evaluate a program of instruction in English as a Second Language (ESL) reading based on schema theory is discussed. The program is used in public schools in Rio de Janeiro (Brazil). In the first section, the research design, procedures, and instruments are presented. An experimental method is used, but is adapted so that ethnographic, qualitative, and process data can be used to interpret quantitative information. The rationale for this approach is discussed. In the second section, the findings of the study are analyzed. In the conclusion, an argument is made for the integration of experimental or quantitative data and ethnographic or quantitative data. This position is supported by the argument that the two kinds of data are of different natures, each with inherent limitations, and thorough research requires the use of quantitative information to clarify or question experimental findings. In this study, it is proposed that the non-experimental data were particularly helpful in calling attention to factors affecting the experiment's internal validity and in enlarging its external validity. A brief bibliography is appended. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED324950

THE EVALUATION OF AN EFL READING SYLLABUS MODEL: INTEGRATING EXPERIMENTAL AND ETHNOGRAPHIC/QUALITATIVE DATA*

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
* Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Luiz Paulo da Moita Lopes (UFRJ)

This paper centers on the discussion of the research paradigm utilised in a pilot evaluation of an EFL schema theoretic reading syllabus model (henceforth, the STR syllabus) in use in a school of the public sector in the city of Rio de Janeiro. In the first section of this work, the research design, procedures and instruments are presented. A paradigm which focuses on experimental and non-experimental (ethnographic/qualitative) data is argued for. In the second, the findings of the investigation are analysed. The paper concludes by demonstrating the need of accounting for ethnographic/qualitative data when experimental research is carried out in the classroom.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Lopes
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

1. The research design

1.1. Justifying this research

Although the STR syllabus model has been designed in a principled manner (cf. Moita Lopes, 1986 and 1987), ie, it has taken into account the local conditions of the ELT Brazilian classroom, theoretical and empirical research in the areas of discourse analysis and schema theories of reading comprehension, and theoretical knowledge of the areas of syllabus design and FL learning, the effectiveness of the STR syllabus can only be checked in the real FL classroom. In other words, although a syllabus may be theoretically sound, before it is actualised in the classroom, there is nothing we can say about its efficiency in language teaching terms.

* The research on which this paper is based is funded by CNPq 300194-86.

FL018932

Therefore, it was necessary to think out a research design which could put the syllabus to the test of practice. To do that, I designed some teaching materials for a beginner's course in the light of the syllabus and investigated its practical validity in the classroom.

1.2. The rationale for this research

Determining an appropriate research design constitutes one of the main concerns of anyone who is involved with research. This is particularly true in an area of investigation such as Applied Linguistics due to the lack of any norm of inquiry widely followed in this field. In fact, this seems to be true of the social sciences in general where there is no universally accepted research paradigm (cf. Thiollent , 1986:21). I believe that what the researcher in this area has to do is to be careful in choosing a paradigm which is adequate to the object of his investigation under the existing pragmatic conditions (cf. Moita Lopes, 1988). Such was the attitude that guided this inquiry.

Research in the social sciences has traditionally followed the nomothetic tradition of the natural sciences, ie, it has followed the experimental method of research through which the researcher tries to predict the effects of a particular variable provided that certain variables are under control.

Accordingly, in applied linguistics, evaluation of innovations in language teaching has normally been undertaken along the lines of this experimental paradigm (cf Allen and Davies, 1977; Hatch and Farhady, 1982, inter alia). Basically, this involves the performing of an experiment. The testing of a hypothesis through: a) the manipulation and/or control of variables - which account for the internal validity of the experiment - so that it can be shown that the outcome of the research is "a function of the factor you have selected rather than other factors you haven't controlled" (Hatch and Farhady, 1982:7); and b) the analysis of the independent

and/or interaction effects of independent variables on a dependent variable. These effects are expressed through objective quantifiable data which are analysed in order to show the statistical significance of a specific innovation in language teaching, ie, of the hypothesis being tested. These results are then generalised to other populations provided that certain variables which may affect the external validity of the experiment are also controlled. In other words, we can explain the effects of the innovation because we can predict their cause. And by describing the experiment, ie, making explicit all the variables under control, statistical analyses used etc., the researcher pr[illegible]s other investigators with the necessary instruments for the [illegible]plication and cross validation of his research.

More recently however another kind of research paradigm has been gaining support in the social sciences. Mainly guided by the fact that the object of research in this area is different from that of the natural sciences, several researchers in the social sciences have argued for the need of paradigms other than the positivist one. In other words, it has been pointed out that researchers in the social sciences must seek their own means of investigation. These authors have argued against what they call the artificiality and restricted scope of the experimental paradigm since by controlling extraneous variables to the experiment and by relying only on quantifiable outcome oriented data, this type of research ignores qualitative and process - oriented data, which center on the social context where the investigation took place. Thus, it is believed that the complexity of the social fact makes it questionable to use the same type of research paradigm typical of the natural sciences in the social sciences, for it is impossible to control and predict the laws which regulate human behaviour. Therefore, the quantitative data provided by the experimental paradigm are of doubtful value and consequently the results of any research can only be evaluated subjectively and not on the basis of statistical data. Other kinds of research models, which are qualitative and process-oriented, have been

suggested, for example, illuminative research (cf. Parlett and Hamilton, 1972), participant research (cf. Brandão, 1981), ethnographic research (cf. van Lier, 1988), phenomenological research (cf. Giorgi, 1985); introspective research (Faerch and Kasper, 1987), action research (cf. Nixon, 1981), etc.

Applied linguists are also increasingly arguing for the use of qualitative/process-oriented data instead of the restriction to experimental research models prevalent in their field: "a shift from what we normally consider important (ie, experimental research) to other forms of research (ie, non-experimental research)" (Ochsner, 1979:58).

In relation to foreign language classroom research, the case in point here, more and more emphasis has been given to the ethnographic paradigm (cf. van Lier, 1988) due to its focus on the whole of the classroom context and on the view that participants of the research have of what is going on in the classroom. The ethnographic researcher observes what is happening in the classroom through the eyes of the participants (students and teachers). He interviews them and asks them to take field-notes and keep diaries, etc. so that through the interpretation of these triangulated data, he can come to a conclusion about what happened in the complexity of the classroom. The researcher therefore does not control any variable which may influence the teaching/learning milieu.

The distinction between these two types of research (experimental and non-experimental) is therefore clear. Essentially, it has to do with whether the researcher operates with a standardised reality or not. Experimental research by eliminating extraneous variables, which could interfere with the researcher's explanations of what caused the experimental results, works with a standardised reality. That is, anything that is not relevant to this model of reality is stripped away so that the researcher can provide an explanation to the effects of the innovation. Quantitative data are then used to <u>explain</u> what happened. On the other hand, non-experimental research involves no standardisation of reality. In fact, this paradigm implies a way of getting closer to reality

as it is, ie, the contextualisation of the research is its main characteristic. And the researcher wants then to understand how a particular innovation functions amidst the complexities of reality rather than explain its effects by standardising reality. Therefore, the more subjective type of data (qualitative and process-oriented data) ignored by an experimental evaluation is made use of since it is more sensitive to the educational context. Also, while the experimental type of research focuses on confirming or not a given hypothesis, ie, it has a retrospective character, the non-experimental kind has a prospective nature in the sense that it allows for continuing improvements of the innovation in the actual process of its application.

Now that the two research traditions have been made clear, it is necessary to discuss the kind of paradigm which oriented the research being reported on here. In accordance with applied linguists such as Brumfit (1984b), Allen (1984), Ochsner (1979), Long (1980 and 1984), inter allia, I took a dualist perspective of research, ie, I tried to provide for quantitative and qualitative data. That is, rather than following one line of investigation and disregarding the other, the research to be discussed here tries to approach the evaluation of the syllabus both from an experimental and from a non-experimental perspective. Each of these approaches provides data of a different nature which can be equally useful in the task of evaluation, making it more thorough by providing different sorts of data which can then be played against each other in the interpretation of results. In this sense, instead of restricting this research to a quantitative paradigm, which provides for objective data in the context of an experiment, related to the outcome of the innovation (product-oriented data), the type of research to be undertaken here is supplemented by an ethnographic/ qualitative approach, which puts forward data connected with the actual process of the teaching/learning experience and which draws on the participants' (teachers and students) view of this process, rather than on the observer's.

I also find support for this rationale in the areas of

behavioural research: "Ideally, we should, whenever possible, approach research problems and test research hypotheses both experimentally and non-experimentally" (Kerlinger, 1973:346); and in the area of evaluation of educational and social programmes:

> "Enthusiasts have championed two competing styles of inquiry: one requires pointed questions and strong controls while the other is naturalistic, broad and exploratory. These two styles are not mutually exclusive and should be combined in various proportions for various purposes" (Cronbach, 1982:x).

The desirability of bringing these two perspectives of research together then has to do with the fact that each one of them gives us something different about the nature of the fact we are examining, ie, two different sorts of insights into the innovation are put forward. While the experimental paradigm decontextualises the research and narrows its focus on the outcome of the innovation, in terms of learners' performance, the non-experimental paradigm contextualises the investigation and broadens its scope by considering variables related to the actual process of learning and teaching and to the participants' perception of this process. In this way, by bringing these two different research perspectives together, an attempt is made to account for their intrinsic limitations. Or, as Cronbach (1982:45) puts it,

> "Something is gained when an evaluation becomes more objective, more reproducible, more concentrated... Something else is gained when the evaluation becomes more phenomenological, more flexible, broader in its coverage".

Furthermore, due to the pedagogic nature of this inquiry, it is crucial that this research make use of these two types of data so that an insight may be gained not only in terms of the innovation as regards learners' performance but also in terms of the conditions within which investigation in the classroom has to take place.

In line with this sort of attitude towards research, the difficulty the investigator faces, I think, has to do with the ability to alternate between the two types of paradigm, since both

styles of research are derived from different views of the research act, which therefore involve different instruments and procedures. Crowle (1976:163, as quoted in Ochsner, 1979:64) has argued that despite the differences in both approaches to research, it is "possible with careful gear changing to alternate between them". The central issue in my view however has to do with reconciling a kind of research paradigm which is exploratory (non-experimental) with one which has clear research mechanisms predetermined before the research act itself (experimental) and which involves hypothesis testing. To my mind, if the researcher agrees with the position taken in this work that it is necessary to have access to both types of data, he has to decide on the paradigm which provides for the kind of data he is primarily interested in and which better fits the circumstances in which he is going to carry out his research, and then somehow try to account for the other type of data which the chosen paradigm does not tap. In other words, the investigator will operate according to the principles of a particular research paradigm, ie, his primary research focus, but at the same time through the use of specific research instruments he will try to consider a different sort of data, ie, his secondary research focus. Therefore, his primary research focus sometimes becomes peripheral and allows for the provision of secondary data. In this way, his research design will make it possible for him to have access to both quantitative/outcome-oriented data and ethnographic/ qualitative/process-oriented data concurrently, which he can then use simultaneously in the interpretation of the results.

In accordance with this position towards research, the design adopted in this work follows an experimental paradigm, but it also tries to account for non-experimental data as secondary data. Despite acknowledging the difficulties in undertaking an experiment when dealing with social behaviour, which is an aspect to be taken into consideration when the experimental results are analysed here, I still found the two following reasons for choosing an experimental paradigm as the main research focus: a) Although I agree that process-oriented data are crucial in making the evaluation more

complete by being more phenomenological, in the present research I was particularly interested in investigating the effectiveness of the STR syllabus approach to the teaching of reading in terms of learners' performance (ie, product-oriented data). Therefore, if I focused on process-oriented data, product-oriented data would be made less relevant since the research would not involve hypothesis testing, control groups etc. b) To make use of a fully ethnographic/qualitative type of evaluation, I would have needed more time than I had to carry out this empirical work. It would have been necessary to spend some time exploring the school environment, before the materials actually started being used, in order to decide on the issues to be inquired into. Further, more time would also have been needed to train teachers in undertaking a fully qualitative type of evaluation.

Therefore, the paradigm used here is experimental on grounds related to the nature of the data particularly relevant to this inquiry and also to pragmatic constraints. However, this research also tries to account for data of an ethnographic/qualitative and process-oriented type in the light of which I will try to further interpret the quantitative findings. Normally, the interpretation of experimental findings rests solely on speculation; yet, here I will make use of non-experimental data to justify my speculations. As I will discuss in 3 below, the non-experimental data will also provide means of checking the external and internal validity of the experiment by putting forward information on whether teachers really followed the treatments assigned to their groups, their biases, etc.

It should also be noted, as I will further discuss in 3, that this research design, by centering on experimental data which are in their turn illuminated by non-experimental data, indicates that experimental findings will be used here not only for their own sake but also as a means of pointing to directions for further investigation.

In the next section, I summarise the research procedures to be utilised in the evaluation of the syllabus.

1.3. The research design of the syllabus evaluation

1.3.1. Experimental research data procedures

Setting up an experiment in the classroom brings about specific difficulties for the experimental researcher due to the impossibility of controlling all the extraneous variables occurring in the classroom. That is, it is not feasible to show that the findings of the research are a result of the independent variable(s) chosen and not of other uncontrolled variables. Nevertheless, despite the difficulty of putting together a true experiment in the classroom, researchers still find cause to make use of a quasi-experimental design (cf. Campbell and Stanley, 1966; Hatch and Farahdy, 1982; and Kerlinger, 1973) because "even a small experiment can be of great value in sharpening observation and developing intuitions" (Clark, 1977:141). What is essential is that the investigator make clear the variables his experiment does not control and take them into consideration when interpreting his results so that the probability that his results were due to uncontrolled factors is assessed.

On the other hand, quite apart from the desirability of undertaking research under very narrow control constraints for the sake of the internal validity of the experiment, when conducting research in the classroom, one also has to consider its pedagogic purpose. If the typical features of the "learning milieu" are stripped away, the classroom loses its representativeness and the external validity of the research and its pedagogic concern are in fact affected. What seems to be essential then is a situation in which there is a balance between the internal and the external validity factors. In this sense, a quasi-experimental research design seems to be a quite adequate alternative for experimental research in the classroom context.

1.3.1.1. Summary of the experimental research data procedures

- Quasi-experimental design
- Duration of experiment: 10 weeks
- School context: 4 schools of the public sector in the city of Rio de Janeiro.
- Number of groups: 4 groups (2 control and 2 experimental groups)
- Schools and teachers:

Group 1: experimental group
Teacher: Miriam Nunes
School: Escola Municipal Francisco Cabrita
Address: Rua Melo Matos, 20 - Tijuca

Group 2: experimental group
Teacher: Sandra Bedran
School: Escola Municipal Carlos Chagas
Address: Rua Romero Zander, 420 - Bonsucesso

Group 3: control group
Teacher: Sergio Pontes da Silva
School: Escola Municipal Maria Izabel Bivar
Address: Rua José dos Reis, 1305 - E. de Dentro

Group 4: control group
Teacher: Regina Gaya da Penha Valle
School: Escola Municipal João Kopke

- Number of subjects: 133 pupils
- Independent variables:

A - treatment - experimental (materials which realize the STR syllabus)
- control (materials reflecting the principles of the "reading method" - cf. Rivers, 1972: 22-24; based on a view of reading as text-based; the same texts of the experimental treatment were used to guarantee the use of the same systemic and schematic elements in the control materials)

B - Teachers' expectations: experimental and control (the inclusion of this independent variable was due to the need to control for the reactive effects of teachers' expectations towards their groups since they were aware that they were taking part in an experiment)

- Dependent variable: subjects' achievement scores on two reading comprehension post-tests

. Post-tests:

Description: Two 20 item reading comprehension tests: one test corresponding to the experimental group treatment (the EXPT, which centers on global comprehension and on process-measures of text comprehension) and the other to the control group treatment (the CONPT, which focuses on the comprehension of every item of the text and on traditional product-measures of text comprehension)

. Test Administration: the 4 groups were required to take both types of post-tests so that the possible backwash effect of a group's specific treatment on the pupils' score on the post-tests was also taken into consideration.

. Pilot work of the post-tests:

Subjects: 55 subjects of the third form of the school of group 1 used as test groups for the experimental group test.

53 subjects of the third form of the school of group 4 used as test groups for control group tests.

Tests checked for reliability: internal consistency of the items (Kuder-Richardson-20)

Reliability coefficient: experimental test - .52
control test - .7

Tests checked for face, content, and construct validity.

- Internal validity: a) checked in terms of the subjects' social background, general learning abilities, sex, age groups, material previously covered. Only problem: experimental groups more used to reading in Portuguese. This difficulty was detected post-facto.

  b) checked in terms of teachers' academic qualifications.

  Pre-Test: 20 items selected from a standardised vocabulary test - The vocabulary in Context Examination, E.L.I. English Achievement Series, by A. Palmer, R. Trasher and J. Upshur. English Language Institute, University of Michigan, 1967.

  Statistical treatment: one-way ANOVA; significance level: .05.

  Objective: to check teachers' claim that learners were virtually zero-beginners.

  Results: the variability of the mean scores of the groups was not significant at the .05.

- External validity: controlled in terms of the Hawthorne and expectancy effects in order to eliminate the reactive effects of testing and of experimental arrangements.

- Distribution of independent variables across the groups:

GROUP 1	GROUP 2
. experimental treatment . teacher had experimental group expectations	. experimental treatment . teacher was led to believe that her group was a control group (no expectations for success).
GROUP 3	**GROUP 4**
. control treatment . teacher had control group expectations	. control treatment . teacher was led to believe that her group was an experimental group (expectations for success)

- Statistical treatment: factorial analysis of variance (two-way ANOVA)
- Objective: to test the significance of the independent and interaction effects of the two independent variables on the dependent variable.
- Research hypothesis: the experimental group subjects will achieve higher scores on both types of post-tests.

  the level of statistical significance pre-established for the ANOVA is .05.

- Factorial design of the experiment

Independent variable A: treatment

Independent variable B: teachers' expectations		A_1 Experimental group	A_2 Control group
B_1	Exp/al group expectations	Group 1 $A_1 B_1$	Group 4 $A_2 B_1$
B_2	Control group expectations	Group 2 $A_1 B_2$	Group 3 $A_2 B_2$

Dependent variable: reading comprehension achievement test scores

1.3.2. Ethnographic/Qualitative research data procedures

The procedures discussed in 1.3.1.1. above have been concerned with the experimental findings of this evaluation, that is, issues connected with the outcome of the implementation of the STR syllabus model in terms of learning effectiveness as expressed in the achievement post-test scores. The focus there was on product-oriented data. In the present section I want to address non-experimental qualitative aspects of this research design, which try

to provide for a subjective type of evaluation as represented in the view of the participants in this research. This type of data also aims at putting forward information about the actual language teaching process that went on in the classroom - process-oriented data - so that the external and internal validity of the experiment described above in 1.3.1.1. may be further checked. As pointed out above in 1.2., , where I discuss the rationale for the type of research design employed in this work, ethnographic/qualitative data here are secondary and are used to illuminate the experimental findings. It is true nevertheless that non-experimental data will be paramount here in projecting the findings of this work into future research (cf. 3 below).

The qualitative process-oriented type of evaluation undertaken here makes use of two research instruments: teachers' diaries and teachers' final evaluation report.

1.3.2.1. Summary of the ethnographic/qualitative research data procedures

- Two research instruments: A: Teachers' diaries
  B: Teachers' final evaluation reports

A - Teachers' diaries

. Objective: to provide for the teachers' view of what went on in the classroom and for access to the actual teaching process in the experimental and control groups.

This kind of data by putting forward information on the actual language teaching process also offers means for checking the internal and external validity of the experiment.

. Description: teachers were asked to record three specific kinds of information at least for each lesson: a) the daily presentation of the materials; b) their attitude to the materials; and c) the pupils' attitude to the materials.

. Pilot work: the instrument was piloted with the four teachers before the experiment started.

B - Teachers' final evaluation reports

. Objective: to have teachers look back at the teaching process as a whole and provide for a final evaluation of what went on in the classroom.

. Description: teachers were asked to tackle three points at least, namely, their views of: a) the syllabus they used in the reading course, as expressed in the materials, in comparison with existing syllabuses; b) the materials of the reading course in comparison with materials previously used by the pupils; c) the pupils' general attitude towards the materials and reading course.

- Research tactics:

The data put forward by these two research instruments were submitted to a hermeneutic mode of inquiry. The meaning of these data is solely dependent on the researcher's interpretation of the teachers' views of what went on in the classrooms. To analyse these subjective findings, one should bear in mind the sorts of biases with which this researcher/reader approaches the teachers' texts. However, this position towards this part of the empirical research is not in fact essentially different from my attitude towards experimental data: experimental data should also be considered in the light of the researcher's biases, research procedures, etc. That is, both types of inquiry (experimental and non-experimental) are to be considered under the same sort of critical attitude towards research: the research act is unavoidably subjective.

In the hermeneutic type of inquiry carried out here, rather than looking for the cause of an event as in the experimental part of this evaluation, I tried to <u>understand</u> what happened by imposing purpose on the event (cf Ochsner, 1979:68). That is, I interpreted the data provided by the diaries and reports by assuming the teachers' purpose in using a particular set of materials in the

classroom. Or, as Ochsner (1979:69) puts it, "By assuming the teleological end, a hermeneutic researcher can interpret an event". My task therefore was to check if the attributed teleological purposes for each teacher in each group (ie, treatment and expectations) held.

2. The findings of this research

This section presents the main findings of this research: the experimental results and the ethnographic/qualitative findings are discussed separately. In the concluding section of this paper (cf. 3 below), these two types of data are related to each other.

2.1. Experimental research findings

2.1.1. Statistical analysis of the CONPT scores

Table I - Breakdown statistics of the CONPT scores

GROUP	MEAN	STD DEV	N
1	9.4000	1.9582	(30)
2	10.5152	2.0019	(33)
3	8.7419	2.0325	(31)
4	9.1714	2.6511	(35)
TOTAL	9.4651	2.2675	(129)

Table II - ANOVA for the means of the groups on the CONPT

Source of variation	SS	df	MS	F	SIGNIF. OF F.
A. Main effects	36.574	2	18.287	3.795	0.025
TREATMENT	31.934	1	31.934	6.627	0.011
EXPECTATIONS	3.405	1	3.405	0.707	0.402
B. Two-way Interactions	19.169	1	19.169	3.978	0.048
TREATMENT by EXPECTATION	19.169	1	19.169	3.978	0.048
EXPLAINED	55.744	3	18.581	3.856	0.011
RESIDUAL	602.349	125	4.819		
TOTAL	658.093	128	5.141		

129 cases were processed
0 cases (0.0 pct) were missing

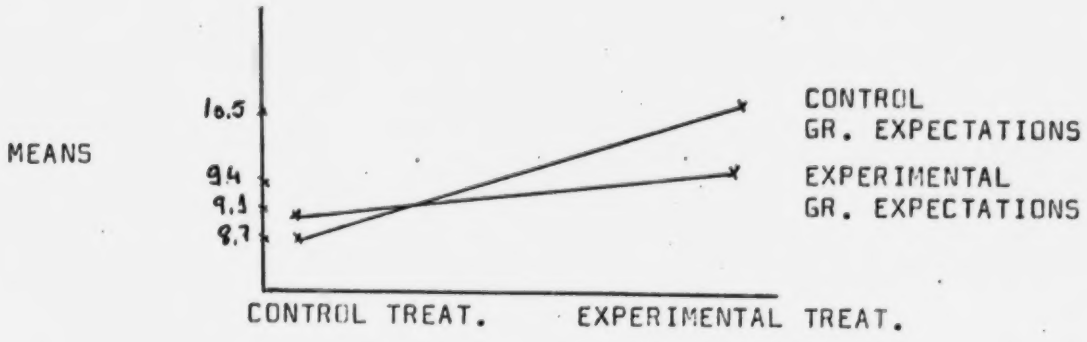
The ANOVA (cf Table II above) has shown that the existing variability between the means of the groups on the CONPT is statistically significant as regards the independent effect of treatment (cf $F(1,125) = 6.627$, $p=0.011$), but it is not significant as concerns the independent effect of expectation (cf $F(1,125) = 0.707$, $p=0.402$). The ANOVA therefore indicates the significance of the main effect of treatment. However, the interaction effect of these two independent variables (treatment and expectation) on the dependent variable (ie, the CONPT) is also significant (cf $F(1,125)= 3.978$, $p=0.048$).

The results of this ANOVA indicate therefore that whereas the independent effect of treatment is significant and of expectation is not significant, the interaction effect between these two independent variables on the dependent variable CONPT is significant.

In other words, the interaction effect between these two independent variables is influencing the results. Figure 1 plots the means of the groups for main effects and interaction effect and

graphically demonstrates both that the experimental groups did better on the CONPT and that there is an interaction effect.

Figure 1 Main effects and interaction effect (CONPT means)



2.1.2. Statistical analysis of the EXPT scores

Table III - Breakdown statistics of the EXPT scores

GROUP	MEAN	STD DEV	N
1	11.9667	2.3995	(30)
2	12.3939	3.1518	(33)
3	9.2258	3.3735	(31)
4	10.8857	2.6654	(35)
TOTAL	11.1240	3.1325	(129)

TABLE IV - ANOVA for the means of the groups on the EXPT

Source of variation	SS	df	MS	F	Signif. of F.
A. Main effects	153.208	2	76.604	8.967	0.000
TREATMENT	144.311	1	144.311	16.893	0.000
EXPECTATIONS	13.164	1	13.164	1.541	0.217
B. Two-way Interactions	35.000	1	35.000	4.097	0.045
TREATMENT by EXPECTATIONS	35.000	1	35.000	4.097	0.045
EXPLAINED	188.208	3	62.736	7.344	0.000
RESIDUAL	1067.808	125	8.542		
TOTAL	1256.016	128	9.813		

129 cases were processed
0 cases (0.0 pct) were missing

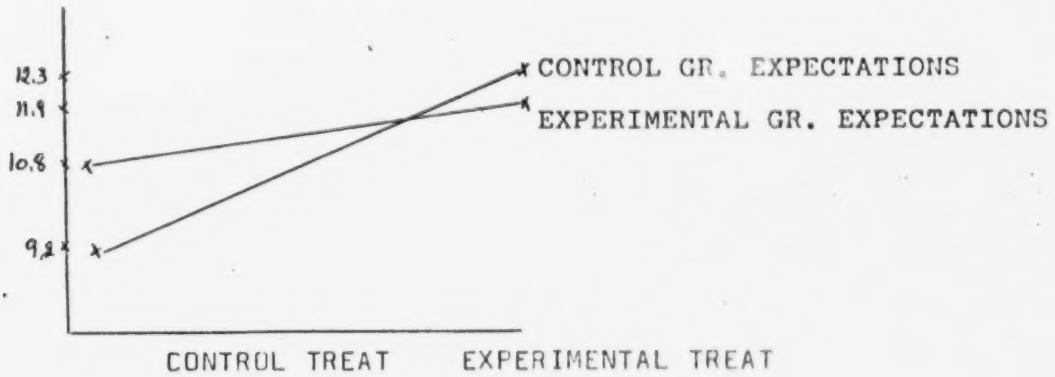
The ANOVA (cf Table IV above) has demonstrated that the existing variability between the means of the groups on the EXPT is statistically significant in terms of the independent effect of treatment (cf $F(1,125= 16.893$, $p= 0.000$), but it is not significant as regards the independent effect of expectation (cf $F(1,125) = 1.541$, $p= 0.217$). That is, the ANOVA again indicates the significance of the main effect of treatment. Nevertheless, the interaction effect of these two independent variables (treatment and expectation) on the dependent variable EXPT is significant (cf $F(1,125) = 4.097$, $p = 0.045$). Note however that the main effect of treatment ($p = 0.000$) is significant way beyond the effect of interaction ($p = 0.045$).

The results of this ANOVA therefore have also shown that whereas the independent effect of treatment is significant and of expectation is not, the interaction effect between these two independent variables on the dependent variable EXPT is significant. In other words, exactly the same pattern of results detected in the ANOVA of the CONPT mean scores is found here. The groups'

performance is consistent on both types of tests. That is, in both cases the interaction effect between the two independent variables may be affecting the results.

Figure 2 below plots the means of the groups for main effects and interaction effect and again graphically indicates not only that the experimental groups did better on the EXPT but also that there was an interaction effect. However, if one compares Figure 2 with Figure 1, one sees that whereas in Figure 1, where the CONPT means are plotted, the larger distance between the scores is between the experimental group scores, in Figure 2 below, where the EXPT means are marked, the larger distance between the scores is between the control group scores. I will come back to this issue in 2.2. below, where I discuss the experimental results.

Figure 2 - Main effects and interaction effect (EXPT means)



2.2. Interpretation of the experimental research findings

As the results of the ANOVAS of the groups' mean scores on both types of post-tests (CONPT and EXPT) have shown, the research hypothesis of this experiment has been confirmed since the experimental groups had higher scores on both post-tests and the variability of the groups' mean scores on both tests is significant at the 0.5 level, as previously established. That is, the independent effect of the treatment variable, as hypothesized, is

statistically significant and the independent effect of the expectation variable is not statistically significant. Yet, the ANOVAs have indicated that the interaction effect between the two independent variables is significant.

This interaction of expectation with treatment however seems to be detected within specific treatment groups when the test type does not correspond to these groups' treatment. That is, when the test is EXPT, the expectation variable will interact with the control treatment, appearing therefore to cause the interaction effect. Conversely, when the test is the CONPT, the expectation variable will interact with the experimental treatment, which may be said to cause the interaction effect. Therefore, to my mind, this interaction effect offers no threat to the confirmation of the hypothesis of this experiment since it consistently occurs between groups of a specific treatment and not across the experimental and control treatments. However, on the basis of data provided by the experiment itself, I have no way of explaining why experimental group 2, whose teacher had control group expectations, did significantly better than group 1, whose teacher had experimental group expectations. I am left with the possibility that the experimental treatment withstands teachers' expectations and that the factors which bring about the difference between these two experimental groups are not captured in the experiment. Different sort of data, related to the actual teaching process, ie, process-oriented data (cf 2.3 below), may clarify this issue above.

Despite the limitations of this experiment in terms of internal and external validity (cf. 1.3.1), I still think that there are grounds for saying that the outcome-oriented data of this experimental evaluation have put forward some evidence of the practical effectiveness of the STR syllabus as reflected in the experimental materials: this syllabus seems to embody a more efficient way of teaching reading than traditional pedagogies, as embodied in the control treatment.

The non-experimental type of data to be discussed in the next section may provide further bases for the internal and external validity of the experiment. In 2.3 below I describe and discuss ethnographic/ qualitative data and in 3 quantitative data and ethnographic/qualitative data are related to each other.

2.3. Ethnographic/qualitative research findings

These findings, besides reflecting how the participants (teachers and pupils) in the experiment affectively perceived what went on in the classrooms, also provide information on the actual language teaching process in such a way that the validity of the experiment can be further examined.

2.3.1. Summary of the data provided by teachers' diaries and final evaluation reports.

Besides using the teachers' diaries and reports as sources for a qualitative evaluation of the materials, the hermeneutic mode of inquiry here had the objective of investigating whether the telelogical purposes I attributed to the teaching/learning event in each of the four groups were supported (cf 1.3.2.). As I will show below in this section, there seems to be evidence that the imputed telelogical ends have been upheld. While I acknowledge the subjectivity of this kind of data - which is in fact its strength - and of my interpretation, I want to emphasise that the evidence found in the diaries seems to be confirmed by the final reports.

Group 1 teachers' diaries and report (experimental treatment and experimental group expectations) demonstrate that she seems to have interpreted the experimental materials adequately and made few modifications in the activities, which do not seem to have altered the experimental treatment and which appear to have been offered in the diaries as factual information on what went on in the classroom. Her attitude and her view of the pupils' attitudes towards the materials and reading lessons are positive. She also

indicates her preference for the reading materials and syllabus in comparison with the existing syllabuses and with the syllabuses she used before the reading course (ie, a structural syllabus with emphasis on oral skills). She taught 18 classes.

Group 2 teachers' diaries and report (experimental treatment and control group expectations) are critical of several parts of the materials and she has actually tried to overcome the faults she found with them by improving several activities. It could be said that she made a big effort to better the materials, which she thought were control materials. I do not think however that these changes have essentially altered the approach she had to follow. In fact, her interpretation of the approach as perceived in these two research instruments seems to have been adequate. Despite her criticism she thinks that the materials work and also that pupils' like them. It seems to me that although she appears to have had control group expectations, the innovative character of the experimental treatment and the pupils' positive attitudes have led her to think that the materials were efficient. Also, while group 3 teacher, who used the control materials, had control group expectations and approved of the objective of his materials, was negative towards them, group 2 teacher, who could be characterised in the same way, except for her experimental materials, reacted positively. That is, it seems plausible to suggest that her attitude towards the materials is due to the effects of the experimental treatment. She also prefers the reading syllabus to the syllabuses she has used before. She has indicated however that although she used a structural syllabus and centered on the four language skills before the reading syllabus, she did not emphasise oral skills. She taught 21 classes.

Group 3 teacher (control treatment and control group expectations) is extremely negative about the materials and sees as the only positive aspect the fact that these materials centre on a more restricted objective, in comparison with the four language skill structural syllabus he used before the reading course. This fact however did not alter his view of the materials. Yet, he

indicates that the pupils had positive attitudes towards the materials although he also classifies them as passive. There is no evidence in his diaries and report that he changed any of the activities although he points out that he had to rush through some parts of the materials. He seems to have adequately interpreted the materials. He taught 20 classes.

Group 4 teachers' diaries and report (control treatment and experimental expectations) are very enthusiastic about the materials and clearly contrast with group 3 teachers' attitude since they used the same kind of materials. She modified a few activities but that does not seem to have altered the control treatment. Her changes however do not appear to have been made to correct something that in her view was wrong with the materials but to better suit particular teaching circumstances. Her interpretation of the materials also appears adequate. She also presents evidence that the materials were perceived by the pupils as an innovation although there is evidence that they faced the classes and tests as normal procedures of their school life. She compares the reading syllabus with the existing syllabuses in Brazil and seems to prefer the former. Contrary to group 3 teacher, she has perceived the use of two devices in the materials (cognates and the cultural proximity of the schematic content area of the text) which are not exploited in the control materials methodology. This fact seems to indicate an attempt on her part to find the positive aspects of the materials. She also taught 20 classes.

In brief, I would like to point out that the ethnographic/ qualitative data discussed in this section have shown that all the four teachers seem to agree on the adequacy of a reading syllabus for Brazilian secondary schools although there have been differences in their attitudes to the two approaches utilised. The experimental treatment, which reflects the STR syllabus model evaluated in this work , appears to have resisted the teacher's expectations from an affective perspective since group 2 teacher (experimental group), despite her control group expectations (as indicated in her diaries and report), still thinks positively of her materials. Conversely,

group 3 teacher, who had control materials and control group expectations, is quite negative about his. All the four teachers have indicated a positive attitude towards the materials on the pupils' part.

In 3 below I relate the experimental to the non-experimental findings of this research in the conclusion of this paper.

3. Conclusion: integrating experimental and ethnographic/qualitative data

The integration of experimental and ethnographic/qualitative data in this research design was due to the fact that because these two types of data are of a different nature and have their own inherent limitations (cf. 1.2), only by tapping both kinds of data can one make the research more thorough. The ethnographic/ qualitative data are not captured in the experiment and may help to clarify or question the experimental findings.

The non-experimental data have been particularly insightful in calling attention to factors which may be said to have affected the internal validity of the experiment. For example:

a) Control group 4 teacher's perception of features of the experimental treatment has possibly made the control treatment similar to the experimental treatment. She perceived the use of two devices which facilitate reading comprehension (ie, cognates and the cultural proximity of the content schematic area of the text) in her materials and which are exploited in the experimental treatment. Therefore, group 4 control treatment may have been closer to the experimental treatment than group 3 control treatment. This aspect may then further explain the reason why group 4 achieved higher mean scores than group 3. That is, in addition to group 4 teacher's experimental group expectations, she may have also dealt with features which are typical of the experimental treatment.

b) Evidence that experimental group 2 teacher made a big effort to improve her materials, which she thought were control

materials. That together with her longer time of instruction (21 classes) and the actual experimental treatment of her group may explain why group 2 did significantly better than group 1 on the CONPT.

c) Evidence that the experimental treatment may have resisted teachers' expectations in affective terms. Experimental group 2 teacher thought positively of her materials despite the fact that she seems to have had control expectations. Therefore, it could again be argued that the experimental treatment was not alone in determining group 2's better performance.

In the context of the experimental part of this evaluation, the limitations of the experimental data had been virtually reduced to a variable which was a shortcoming of the experiment: the experimental groups seemed to be more used to reading in Portuguese than the control groups, ie, a factor which affects the internal validity of the experiment (cf. 1.3.1.1). Therefore, the factors brought to light in the non-experimental part of this evaluation would not have been captured had this research been restricted to experimental data. In other words, by making the research ethnographic/qualitative, factors which may question the internal validity of the experiment have been more clearly exposed. In this way, one becomes more aware of the scope of the claims of the experimental findings.

On the other hand, the non-experimental findings seem to have enlarged the external validity of the experiment since the non-experimental data appeared to confirm that the teachers' expectations seemed to have been manipulated and no reactive effects towards the tests and the experimental arrangement on the part of the pupils were noted. Yet, because of constraints related to the size of the sample used here, to the short period of time this experiment involved, and in fact because of issues which affected the internal validity of the experiment, generalisations from this experiment to other EFL teaching situations are to be made with caution.

Nevertheless, while acknowledging that there are points which question the experimental findings above and particularly in view of issues brought about in the non-experimental part of the research, I would like to suggest that the experimental results seem to indicate a trend in the expected direction, ie, the practical effectiveness of the syllabus model - as reflected in the experimental materials - if compared with traditional pedagogies for the teaching of reading. This trend seems to reveal that it is worth pursuing further inquiry through the actual implementation of the STR syllabus.

Furthermore, there is a meta-research point which this inquiry has made clear. This investigation has empirically demonstrated that experimental research in the classroom context should only be done if ethnographic/qualitative data are taken into account otherwise it is impossible to understand what experimental findings in the classroom really mean.

The research paradigm used here by relating experimental data to non-experimental data has shown that:

a) experimental findings - which would have been considered definitive had this research been solely experimental - have been further clarified or become questionable;

b) without the use of ethnographic/qualitative data, my interpretation of the experimental findings would have been simply speculative; and

c) the ethnographic/qualitative data were useful in projecting the findings of this paper into future research by bringing about different factors of importance which need to be further investigated in the actual implementation of the STR syllabus in the classroom, for example: the different rhythm of instruction used by different teachers, control group 4 teachers' perception of features of the experimental methodology, the pupils' positive attitude towards the materials in the 4 groups as a reflection of their perception of the materials as an innovation, etc.

In short, in experimental classroom research it is useful to rely on ethnographic/qualitative data because they make experimental findings less trivial and because, due to their prospective nature, they project the findings of retrospective research, such as the one discussed here, into future research.

BIBLIOGRAPHY

ALLEN, J.P.B. (1984). "General - Purpose Language Teaching: a Variable Focus Approach". In Brumfit, C.J. (ed.), (1984a), pp. 61-74.

ALLEN, J.P.B. and DAVIES, A. (ed .)(1977). Testing and Experimental Methods. The ECAL. vol. 3. London: Oxford University Press

BRANDÃO, C. R. (org.) (1981). Pesquisa Participante. São Paulo: Brasiliense.

BRUMFIT, C.J. (1984a). General English Syllabus Design: Curriculum and Syllabus Design for the General English Classroom. ELT Documents, 118. Oxford: Pergamon.

BRUMFIT, C. J. (1984b). "Function and Structure of a State School Syllabus for Learners of Second and Foreign Languages with Heterogeneous Needs". In Brumfit, C. J. (ed.) (1984a), pp. 75-82.

CAMPBELL, D. T. and STANLEY, J. C. (1966). Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally College Publishing Company.

CLARK, R. (1977), "The Design and Interpretation of Experiments". In Allen, J. P. B. and Davies, A. (eds.) (1977), pp.105-145.

CRONBACH, L. J. (1982). Designing Evaluations of Educational and Social Programs. San Francisco: Jossey - Bass Inc. Publishers.

FAERCH, C. and KASPER, G. (1987). Introspection in Second Language Research. Clevedon: Multilingual Matters Ltd.

GIORGI, A. (ed.) (1985). Phenomenology and Psychological Research. Pittsburgh: Duquesne University Press.

HATCH, E. and FARHADY, H. (1982). Research Design and Statistics for Applied Linguistics. Rowley, Mass.: Newbury House.

KERLINGER, F.N. (1973). Foundations of Behavioural Research. New York: Holt, Rinehart and Winston.

LONG, M. (1980). "Inside the 'Black Box': Methodological Issues in Classroom Research on Language Learning". Language Learning, vol. 30, no.1, pp. 1-42.

LONG, M. (1984). "Process and Product in ESL Program Evaluation". TESOL Quarterly, vol. 18, no. 3, pp. 409-425.

MOITA LOPES, L.P. da (1986). Discourse Analysis and Syllabus Design: an Approach to the Teaching of Reading. Tese de Doutoramento. Universidade de Londres.

MOITA LOPES, L. P. da (1987). "Elaboração de Programas de Ensino de Línguas Estrangeiras: Um Modelo Operacional". Revista Perspectiva, no. 8, pp. 40-51.

MOITA LOPES, L. P. da (1988). "Repensando Métodos de Investigação na Sala de Aula de Inglês: entre o Produto e o Processo de Ensino/Aprendizagem". Conferência proferida no VIII ENPULI, UNB, mimeo,

NIXON, J. (ed.) (1981). A Teacher's Guide to Action-Research. London: Grant Mc Intire Ltd.

OCHSNER, R. (1979). "A Poetics of Second-Language Acquisition". Language Learning, vol. 29, no.1, pp. 53-80.

PARLETT, M. and HAMILTON, D. (1972). Evaluation as Illumination: a New Approach to the Study of Innovation Programs. Occasional Paper 9. Centre for Research in the Educational Sciences. University of Edinburgh.

RIVERS, W. M. (1972). Teaching Foreign-Language Skills. Chicago: The University of Chicago Press.

THIOLLENT, M. (1986). Metodologia da Pesquisa - Ação. São Paulo: Cortez.

VAN LIER, L. (1988). The Classroom and the Language Teacher. Londres: Longman.

END

U.S. Dept. of Education

Office of Education
Research and
Improvement (OERI)

ERIC

Date Filmed

March 29, 1991